

AD-A105 659

NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS
INFORMATION CAPACITY OF GAUSSIAN CHANNELS.(U)
SEP 81 C R BAKER

F/G 9/4

UNCLASSIFIED

N00014-75-C-0491

NL

[OF]
[20]
[20]

[20]

END

DATE

FILED

11 11

DTIC

① LEVEL II

AD A105659

INFORMATION CAPACITY OF GAUSSIAN CHANNELS

Charles R. Baker*
Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

DTIC
ELECTE
OCT 16 1981
S D
B

DTIC FILE COPY

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

*Research supported by ONR Contract N00014-75C-0491 ✓

81 10 15

Abstract

Consider a communication channel without feedback, with transmitted signal $A(X)$, and with additive Gaussian noise N . The information capacity of this channel is obtained subject to the constraints $E||A(X)||_W^2 \leq P$, where $||\cdot||_W$ can be regarded as the RKHS norm of the stochastic process W . The class of admissible processes W includes all Gaussian processes having induced measure equivalent to the measure induced by N .

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

Introduction

This paper considers the information capacity of the Gaussian channel without feedback. The channel input is a sample function from a stochastic process X ; it is encoded into the signal sample function by a coding operation A ; the channel adds a sample function from a Gaussian noise process N (independent of X); the channel output is then a sample function from the process $Y = A(X) + N$. The quantity of interest is the information capacity

$$C = \sup_Q I[X, A(X) + N]$$

where $I[U, V]$ is the mutual information of the processes U and V , and Q is a set of (A, X) defined by appropriate constraints.

Let $\|\cdot\|_N$ denote the reproducing kernel Hilbert space (RKHS) of N ; under the constraint $E\|A(X)\|_N^2 \leq P$, a complete solution to the information capacity problem is given in [1]. However, there is also considerable interest in the capacity problem using constraints of a different type. The following constraint has been examined in special cases in various publications [2], [3], [4], [6]: Q is the set of all (A, X) such that $E\|A(X)\|_W^2 \leq P$, where W is a second Gaussian process. There are various motivations for this definition of Q . For example, if W is the Wiener process on $[0, T]$, as in [2], and $[A(X)]_t = \int_0^t V_s ds$, then $\|A(X)\|_W^2 = \int_0^T V_t^2 dt$. This is the usual power constraint. A more general motivation is that in many applications one may not know the precise covariance of N . It is then of interest to calculate upper and lower bounds for the capacity, over all Gaussian processes N whose induced measures are mutually absolutely continuous to the measure induced by some "most likely" (or reference) Gaussian process W .

A complete solution is obtained here for the capacity problem using the above definition of Q . In a subsequent publication, the results obtained here will be applied in an analysis of the feedback channel of [2].

Mathematical Structure

The channel model is defined as in [1]. The channel noise N is represented by a measure μ_N on the Borel σ -field of a real separable Hilbert space H_2 . The message X is modeled by a measure μ_X on a real separable Hilbert space H_1 . The inner product on H_1 is denoted by $\langle \cdot, \cdot \rangle_1$. $A: H_1 \rightarrow H_2$ is a Borel-measurable coding function. μ_W is a strong second order measure on H_2 ($\int_{H_2} \|x\|_2^2 d\mu_W(x) < \infty$). R_N and R_W denote the covariance operators of μ_N and μ_W . The following assumptions are made.

- (1) $R_N = R_W^{1/2}(I+S)R_W^{1/2}$, where S is a compact operator in H_2 , and S does not have -1 as an eigenvalue. Without loss of generality, it is assumed that $\overline{\text{range}(R_N)} = H_2$.

- (2) The admissible set Q is the set of all (A, μ_X) such that

$$\int_{H_1} \|R_W^{-1/2} A(x)\|_2^2 d\mu_X(x) \leq P.$$

The definition of Q given in (2) implies that $\mu_X \circ A^{-1}$ is strong second order. Since we can (and do) assume that μ_N , μ_W , and $\mu_X \circ A^{-1}$ each have zero mean, the constraint in (2) is equivalent to $\text{trace } R_W^{-1/2} R_{A(X)} R_W^{-1/2} \leq P$, where $R_{A(X)}$ is the covariance operator of $A(X)$. Note that $R_W^{-1/2}$ exists because of assumption (1). If $\overline{\text{range}(R_N)} \neq H_2$ in a given problem, one can replace H_2 WLOG by $\overline{\text{range}(R_N)}$. Alternatively, one could use the original H_2 and apply the constraint

$$\int_{H_2} \sum_n \delta_n^{-1} \langle A(x), Z_n \rangle_2^2 d\mu_X(x) \leq P, \text{ where } R_W = \sum_n \delta_n Z_n \otimes Z_n, \text{ with each } \delta_n > 0 \text{ and } \{Z_n, n \geq 1\} \text{ an o.n. set.}$$

Much of the following analysis will hinge on the properties of the strictly negative eigenvalues of S . These eigenvalues will be designated as $\{\lambda_n, n \geq 1\}$,

$\lambda_n \leq \lambda_{n+1}$, with $\{e_n, n \geq 1\}$ associated o.n. eigenvectors. Of course, depending on the particular R_N and R_W , the set $\{\lambda_n, n \geq 1\}$ can be empty, finite, or countably infinite.

The joint measure on $H_1 \times H_2$ representing message and channel output is μ_{XY} , with μ_Y the measure representing the channel output process Y . They are defined as in [1]:

$$\mu_Y(B) = \mu_X \otimes \mu_N\{(x,y): A(x) + y \in B\}$$

$$\mu_{XY}(C) = \mu_X \otimes \mu_N\{(x,y): (x, A(x)+y) \in C\}$$

where $\mu_X \otimes \mu_N$ is product measure on $H_1 \times H_2$ (all measures are defined on the usual Borel σ -fields, as in [1]). With these definitions, $I[X, A(X) + N] = I[\mu_{XY}]$, where $I[\mu_{XY}] = \infty$ if μ_{XY} is not absolutely continuous with respect to $\mu_X \otimes \mu_Y$, and otherwise

$$I[\mu_{XY}] = \int_{H_1 \times H_2} \log \left[\frac{d\mu_{XY}}{d\mu_X \otimes \mu_Y}(x,y) \right] d\mu_{XY}(x,y).$$

Two results proved in [1] will be central to the following analysis. The first is that for any fixed covariance operator $R_{A(X)}$ of the signal process, the information $I(\mu_{XY})$ is maximized by choosing $A(X)$ to be Gaussian [1; Lemma 6]. The second is that if $A(X)$ is Gaussian with covariance operator

$$R_{A(X)} = \sum_i \alpha_i [R_N^{1/2} v_i] \otimes [R_N^{1/2} v_i]$$

where $\sum \alpha_i < \infty$ and $\{v_n, n \geq 1\}$ is an o.n. set in H_2 , then $I[\mu_{XY}] = (1/2) \sum_n \log[1 + \alpha_n]$ [1, pp. 83-84].

Finally, it is noted that $\|R_W^{-1/2} x\|_2 \equiv \|x\|_W$ can be viewed as the norm of x in the RKHS defined by the kernel $r_W(t,s): H_2 \times H_2 \rightarrow \mathbb{R}$, $r_W(t,s) = \langle R_W t, s \rangle_2$. r_W actually determines a RKHS of real-valued functions on H_2 ; however, this RKHS is a subset of the bounded linear functionals on H_2 , and thus can be regarded as a subset of H_2 .

Main Result

The solution for the information capacity problem defined in the preceding section is given in the following theorem.

Theorem. Let $C = \sup_Q I[\mu_{XY}]$, where Q is defined by assumptions (1) and (2) of the preceding section.

(a) If H_2 is finite-dimensional, then

$$C = \frac{1}{2} \sum_{n=1}^K \log \left[\frac{\sum_{i=1}^K (1+\gamma_i) + P}{K(1+\gamma_n)} \right]$$

where K is the largest integer such that $\sum_1^K \gamma_n + P > K\gamma_K$, and $\gamma_1 \leq \gamma_2 \leq \dots$ is the set of all eigenvalues of S .

(b) If H_2 is infinite-dimensional, $\{\lambda_n, n \geq 1\}$ is not empty, and $\sum_1^K \lambda_n + P > K\lambda_K$ for all negative eigenvalues λ_K of S , then $\sum_n |\lambda_n| \leq P$ and

$$C = \frac{1}{2} \sum_n \log \left[\frac{1}{1+\lambda_n} \right] + 1/2 \left[P + \sum_m \lambda_m \right].$$

(c) If H_2 is infinite-dimensional, $\{\lambda_n, n \geq 1\}$ is not empty, and there exists a largest integer K such that $\sum_1^K \lambda_i + P > K\lambda_K$ and $\lambda_K < \sup\{\lambda_n, n \geq 1\}$ then

$$C = \frac{1}{2} \sum_{n=1}^K \log \left[\frac{P + \sum_{i=1}^K (1+\lambda_i)}{K(1+\lambda_n)} \right]$$

(d) If H_2 is infinite-dimensional, and $\{\lambda_n, n \geq 1\}$ is empty, then $C = P/2$.

In (a) and (c), the capacity can be attained. In (a), it is attained using a Gaussian signal with covariance

$$R_{A(X)} = \sum_{n=1}^K \tau_n [R_N^{1/2} v_i] \otimes [R_N^{1/2} v_i].$$

where $\{v_1, \dots, v_K\}$ is an o.n. set, $Sv_i = \gamma_i v_i$, and $\tau_n = (1 + \gamma_n)^{-1} \left[\frac{\sum_{i=1}^K \gamma_i + P}{K} - \gamma_n \right]$.

The capacity in (c) is attained with a Gaussian signal having covariance of the same form as that for (a), but with γ_i replaced by λ_i and v_i replaced by e_i , $i=1, \dots, K$.

In (b) and (d), the capacity cannot be attained, except in (b) for the special case where $P = -\sum_n \lambda_n$.

In (b) and (c), the capacity is strictly greater than the capacity obtained using the constraint $E_{\mu_X} \|A(X)\|_N^2 \equiv \int_{H_1} \|R_N^{-1/2} A(x)\|_2^2 d\mu_X(x) \leq P$; in (d), the capacity is the same. In (a), the capacity is strictly less than that for the constraint $E_{\mu_X} \|A(X)\|_N^2 \leq P$ if all $\gamma_n \geq 0$ with $\gamma_K > 0$; strictly greater if all $\gamma_n \leq 0$ and $\gamma_1 < 0$; and no general statement holds if $\gamma_K > 0$ and $\gamma_1 < 0$.

Proof of the Main Result

The proof of the Theorem will be given after having obtained several lemmas.

Lemma 1 Let $N \geq 1$ be fixed, and suppose $\{\rho_n, n=1, \dots, N\}$ is a set of strictly positive scalars, $\rho_1 \geq \rho_2 \geq \dots \geq \rho_N$. Let $A \subset \mathbb{R}^N$ be the set of all x such that $x_n \geq 0$ for $n=1, \dots, N$, and $\prod_{n=1}^k x_n \leq \prod_{n=1}^k \rho_n$ for $k=1, \dots, N$. Then $\sup_A \prod_{n=1}^N [1+x_n] = \prod_{n=1}^N [1+\rho_n]$.

Proof Define $G_N: \mathbb{R}^N \rightarrow \mathbb{R}$ by $G_N(x) = \sum_{n=1}^N \log[1+x_n]$. It is sufficient to show that $G_N(x) \leq G_N(\rho)$, subject to the constraints

$$-x_n \leq 0 \quad n=1, \dots, N \quad (C_1)$$

$$\sum_{n=1}^k \log x_n - \sum_{n=1}^k \log \rho_n \leq 0 \quad k=1, \dots, N \quad (C_2).$$

The constraints C_1, C_2 define a convex set in \mathbb{R}^N , and G_N is concave on \mathbb{R}^N .

Thus, one has the problem of maximizing a concave function over a convex set in \mathbb{R}^N ,

with concave and differentiable constraints; by Kuhn-Tucker theory [7] any solution to this problem will define a global maximum over the set A. A solution is any \bar{x}^* in \mathbb{R}^N satisfying the following set of equations for some γ, β in \mathbb{R}^N such that $\gamma_i \leq 0$ and $\beta_i \leq 0$ for $i=1, \dots, N$ [7]:

$$\frac{1}{1+x_n^*} - \gamma_n + \frac{1}{x_n^*} \sum_{k \geq n} \beta_k = 0, \quad n=1, \dots, N \quad (S_1)$$

$$\gamma_n x_n^* = 0 \quad \text{and} \quad -x_n^* \leq 0, \quad n=1, \dots, N \quad (S_2)$$

$$\left. \begin{aligned} \beta_k \left[\sum_{n=1}^k \log x_n^* - \sum_{n=1}^k \log \rho_n \right] &= 0 \\ \sum_{n=1}^k \log x_n^* - \sum_{n=1}^k \log \rho_n &\leq 0 \end{aligned} \right\} \quad k=1, \dots, N \quad (S_3).$$

The system of equations (S_1) , (S_2) , (S_3) is easily seen to be solved by taking

$x_n^* = \rho_n$, $\gamma_n = 0$, and $\sum_{k=n}^N \beta_k = -\rho_n/(1+\rho_n)$ for $n=1, \dots, N$. The fact that this solution gives $\beta_k \leq 0$ for all $k=1, \dots, N$ follows by induction on $N - k+1$, using $\rho_n \leq \rho_{n-1}$ for $2 \leq n \leq N$. \square

Lemma 2 Let $\{x_n^2, n \geq 1\}$ be a summable sequence, with $x_n^2 \geq x_{n+1}^2$. Suppose that S is a compact operator in H_2 which is symmetric, has $M \leq \infty$ strictly negative eigenvalues, and $I + S$ is strictly positive. Then for any finite $N \leq M$ and any o.n. set $\{u_1, \dots, u_N\}$,

$$\prod_{n=1}^N [1+x_n^2] \|(I+S)^{1/2} u_n\|^{-2} \leq \prod_{n=1}^N [1+x_n^2 (1+\lambda_n)^{-1}]$$

where $\lambda_1 \leq \lambda_2 \leq \dots$ are the negative eigenvalues of S .

Proof Using the inequality $\|(I+S)^{1/2} u_n\|_2^{-2} \leq \|(I+S)^{-1/2} u_n\|_2^2$, it suffices to show that $\prod_{n=1}^N [1+x_n^2] \|(I+S)^{-1/2} u_n\|_2^2 \leq \prod_{n=1}^N [1+x_n^2 (1+\lambda_n)^{-1}]$.

Let $\{u_n, n \geq 1\}$ be any fixed CON set in H_2 . Define $X = \sum_n |x_n| u_n \otimes u_n$. The previous inequality will be proved if one shows that

$$\prod_1^N [1 + \|(I+S)^{-1/2} X u_n\|_2^2] \leq \prod_1^N [1 + x_n^2 (1 + \lambda_n)^{-1}].$$

By a result of Horn [5],

$$\prod_1^k \|(I+S)^{-1/2} X u_n\|_2^2 \leq \prod_1^k \tau_n(p) |x_n|^{2-1/4p}$$

for all $k \leq N$ and any fixed $p \geq 1$, where $\{\tau_n(p), p \geq 1\}$ are the eigenvalues of $X^{1/2p} (I+S)^{-1} X^{1/2p}$ and $\tau_n(p) \geq \tau_{n+1}(p)$. This follows from the fact that $(I+S)^{-1/2} X^{1/2p}$ is compact, as is $X^{1-1/2p} = \sum |x_n|^{1-1/2p} u_n \otimes u_n$.

Let $\alpha > 0$ be given. For sufficiently large p , $(1+\alpha)(I+S)^{-1} > X^{1/2p} (I+S)^{-1} X^{1/2p}$ and then $\prod_1^k \tau_n(p) < (1+\alpha)^k \prod_1^k (1+\lambda_n)^{-1}$, so that

$$\prod_1^k \|(I+S)^{-1/2} X u_n\|_2^2 < (1+\alpha)^k \prod_1^k (1+\lambda_n)^{-1} |x_n|^{2-1/4p}, \quad k = 1, 2, \dots, N.$$

Thus, for all $\alpha > 0$, $\prod_1^k \|(I+S)^{-1/2} X u_n\|_2^2 < (1+\alpha)^k \prod_1^k (1+\lambda_n)^{-1} x_n^2$, giving

$$\prod_1^k \|(I+S)^{-1/2} X u_n\|_2^2 \leq \prod_1^k (1+\lambda_n)^{-1} x_n^2, \quad \text{for } k = 1, 2, \dots, N.$$

Applying Lemma 1, this yields

$$\prod_1^N [1 + \|(I+S)^{-1/2} X u_n\|_2^2] \leq \prod_1^N [1 + x_n^2 (1 + \lambda_n)^{-1}]. \quad \square$$

Lemma 3 Suppose that all the eigenvalues $\{\lambda_n, n \geq 1\}$ of S are strictly negative, $\lambda_1 \leq \lambda_2 \leq \dots$, with $S e_i = \lambda_i e_i$, $i \geq 1$, and $\{e_i, i \geq 1\}$ an o.n. set.

(a) If there exists a largest integer $K \geq 1$ such that $\sum_1^K \lambda_n + P > K \lambda_K$, then

$$C = \frac{1}{K} \sum_{n=1}^K \log \left[\frac{\sum_{i=1}^K (1 + \lambda_i) + P}{K(1 + \lambda_n)} \right].$$

The capacity can be achieved; it is attained with a Gaussian signal having covariance operator

$$R_A(X) = \sum_1^K \tau_i [R_N^{1/2} e_i] \otimes [R_N^{1/2} e_i]$$

where $\tau_i = \frac{(1+\lambda_i)^{-1}}{K} \left(\sum_{n=1}^K \lambda_n + P - K\lambda_i \right)$, $i=1, \dots, K$.

This result includes the case when H_2 is finite dimensional (so that $K \leq \dim(H_2)$).

Moreover, this result is also obtained when one adds the additional restriction that support $(\mu_X \circ A^{-1}) \leq M < \infty$. With this restriction, $K = M$ if $\sum_1^M \lambda_n + P > M\lambda_M$; otherwise there exists $K < M$ such that the above expression is the capacity.

(b) If there does not exist a largest integer K such that $\sum_1^K \lambda_n + P > K\lambda_K$,

then $\sum_n |\lambda_n| \leq P$, and

$$C = \frac{1}{2} \sum_{n=1}^{\infty} \log \left[\frac{1}{1+\lambda_n} \right] + \frac{1}{2} (P + \sum_n \lambda_n).$$

Capacity can be attained only if $P = -\sum \lambda_n$; it is then attained by a Gaussian signal with covariance $R_{AX} = \sum_1^{\infty} \tau_i [R_N^{\frac{1}{2}} U^* e_i] \otimes [R_N^{\frac{1}{2}} U^* e_i]$ where $\tau_i = \frac{-\lambda_i}{1+\lambda_i}$. If $P > -\sum \lambda_n$, then the capacity is the limit of the mutual information for a sequence of Gaussian signals (μ_{AX}^M) , with covariance $R_{AX}^M = \sum_1^M \tau_i^M [R_N^{\frac{1}{2}} U^* e_i] \otimes [R_N^{\frac{1}{2}} U^* e_i]$ and

$$\tau_i^M = \frac{(1+\lambda_i)^{-1}}{M} \left(\sum_1^M \lambda_n + P - M\lambda_i \right) \quad i=1, \dots, M.$$

Proof

From the results of [1, pp. 83-84], $I[\mu_{XY}] = \frac{1}{2} \sum_{n=1}^{\infty} \log[1+\tau_n]$ when $\mu_X \circ A^{-1}$ is Gaussian with covariance operator

$$R_{A(X)} = \sum_{n=1}^{\infty} \tau_n [R_N^{1/2} u_n] \otimes [R_N^{1/2} u_n]$$

with $\{u_n, n \geq 1\}$ any o.n. set in H_2 . By Assumption (1), $R_N^{1/2} = R_W^{1/2} (I+S)^{1/2} U^*$ with U unitary. Thus

$$\begin{aligned} E \|R_W^{-1/2} A(X)\|_2^2 &= \text{Trace } R_W^{-1/2} R_{A(X)} R_W^{-1/2} \\ &= \text{Trace } (I+S)^{\frac{1}{2}} U^* R_N^{-\frac{1}{2}} R_{A(X)} R_N^{-\frac{1}{2}} U (I+S)^{\frac{1}{2}} \\ &= \sum_n \tau_n \| (I+S)^{\frac{1}{2}} U^* u_n \|_2^2. \end{aligned}$$

Moreover, for any choice of covariance operator $R_{A(X)}$, the information is maximized if $\mu_X \circ A^{-1}$ is Gaussian [1, Lemma 6]. Thus, one can assume that $\mu_X \circ A^{-1}$ is Gaussian. The capacity problem now reduces to finding $C = \sup_{Q'} \frac{1}{2} \sum_n \log[1 + \tau_n]$, where Q' is the set of all $\{(\tau_n), (U_n)\}$ such that $\tau_n \geq 0$ for $n \geq 1$, $\sum_n \tau_n < \infty$, $\{u_n, n \geq 1\}$ is an o.n. set, and $\sum_n \tau_n \|(I+S)^{\frac{1}{2}} U^* u_n\|_2^2 \leq P$.

We rewrite the preceding expression as

$$C = \sup_{Q'} \sum_n \log[1 + x_n^2 \|(I+S)^{\frac{1}{2}} U^* u_n\|_2^{-2}],$$

where $x_n^2 \equiv \tau_n \|(I+S)^{\frac{1}{2}} U^* u_n\|_2^2$. Since $\{x_n^2, n \geq 1\}$ and $\{\tau_n, n \geq 1\}$ are summable, and $\{u_n, n \geq 1\}$ is to be selected, one can assume WLOG that $x_n^2 \geq x_{n+1}^2, n \geq 1$. From Lemma 2, for any such choice of $\{x_n^2, 1 \leq n \leq M\}$,

$$\sum_1^M \log[1 + x_n^2 \|(I+S)^{\frac{1}{2}} U^* u_n\|_2^{-2}] \leq \sum_1^M \log[1 + x_n^2 (1 + \lambda_n)^{-1}].$$

We will maximize the right side of this inequality for fixed M , and show that the maximum can be attained by $\{x_n^2, 1 \leq n \leq M\}$ such that $\sum_1^M x_n^2 = P$ and $x_n^2 \geq x_{n+1}^2$ for $n=1, \dots, M$.

For fixed $M \geq 1$, define $f_M: \mathbb{R}^M \rightarrow \mathbb{R}$ by $f_M(\chi) = \sum_{n=1}^M \log[1 + y_n (1 + \lambda_n)^{-1}]$. We seek to maximize $f_M(\chi)$ subject to the constraints

$$g(\chi) \equiv \sum_1^M y_n - P \leq 0,$$

$$h_i(\chi) \equiv -y_i \leq 0, \quad i=1, \dots, M.$$

This is a constrained maximization problem. Since $\log(1 + \alpha y)$ is concave over $\{y: y \geq 0\}$ for any $\alpha > 0$, the function f_M is concave over the convex set $\{\chi \in \mathbb{R}^M: \chi_i \geq 0, i=1, \dots, M\}$. Moreover, each constraint function is concave. Thus, any solution to this problem will define a global maximum for f_M [7]. In order for χ^* to be a solution, it is necessary and sufficient that the following set of equations be satisfied [7]:

$$\frac{1}{1+y_i^*+\lambda_i} + \beta - \gamma_i = 0 \quad i=1,\dots,M \quad (1)$$

$$\sum_1^M y_n^* - P \leq 0, \quad \beta [\sum_1^M y_n - P] = 0 \quad (2)$$

$$-y_i^* \leq 0, \quad \gamma_i y_i^* = 0, \quad i=1,\dots,M \quad (3),$$

for some set of non-positive real numbers $\{\beta, \gamma_1, \dots, \gamma_M\}$.

We first attempt to obtain a solution y^* by setting $\gamma_1 = \gamma_2 = \dots = \gamma_M = 0$.

This requires

$$\frac{1}{1+y_i^*+\lambda_i} = -\beta, \quad i=1,\dots,M; \quad \text{thus}$$

$$\sum_1^M y_i^* + \sum_1^M (1+\lambda_i) = -M\beta^{-1}, \quad \text{and}$$

$$y_n^* = \frac{\sum_1^M y_i + \sum_1^M (1+\lambda_i)}{M} - (1+\lambda_n)$$

for $n=1,2,\dots,M$. This definition of y^* and constraints (3) require that

$$\sum_1^M y_i + \sum_1^M (1+\lambda_i) \geq M(1+\lambda_n), \quad n=1,\dots,M;$$

this inequality is satisfied for all $n \leq M$ if and only if it holds for $n = M$. Also

$\beta = -(1+y_i^*+\lambda_i)$ implies $\beta < 0$, so that $\sum_1^M y_i^* = P$ by constraints (2). Thus, if

$\sum_1^M P + \sum_1^M \lambda_i \geq M\lambda_M$, then an optimum solution is given by

$$y_i^* = \frac{P + \sum_1^M \lambda_j - M\lambda_i}{M}, \quad i=1,\dots,M.$$

If there exists $K < M$ such that

$$P + \sum_1^K \lambda_i \geq K\lambda_K$$

$$P + \sum_1^{K+1} \lambda_i < (K+1)\lambda_{K+1},$$

then constraints (1)-(3) are satisfied by choosing

$$\beta = -K[P + K + \sum_1^K \lambda_i]^{-1}$$

$$\gamma_1 = \gamma_2 = \dots = \gamma_K = 0$$

$$\sum_1^K y_i = P$$

$$y_i = 0, i > K$$

$$y_i = K^{-1}[P + \sum_{n=1}^K \lambda_n - K\lambda_i], \quad i \leq K$$

$$\gamma_i = -K[P + K + \sum_1^K \lambda_n]^{-1} + (1+\lambda_i)^{-1} \quad i > K.$$

Thus, under the assumptions of (a),

$$\sup_{Q''} \sum_1^\infty \log[1 + x_n^2(1+\lambda_n)^{-1}] = \sum_{n=1}^K \log \left[\frac{\sum_{i=1}^K \lambda_i + P+K}{K(1+\lambda_n)} \right]$$

where $Q'' = \{(x_n^2): \sum_1^\infty x_n^2 \leq P\}$. As already noted, when $\mu_X \circ A^{-1}$ is Gaussian with covariance operator $R_{A(X)} = \sum_1^\infty \tau_n [R_N^{1/2} u_n] \otimes [R_N^{1/2} u_n]$, $\{u_n, n \geq 1\}$ an o.n. set, then

$$I[\mu_{XY}] = \frac{1}{2} \sum_{n=1}^\infty \log[1 + \tau_n] \leq \frac{1}{2} \sum_{n=1}^\infty \log[1 + x_n^2(1+\lambda_n)^{-1}]$$

if $x_n^2 \equiv \tau_n || (I+S)^{1/2} u_n ||_2^2$, $x_n^2 \geq x_{n+1}^2$.

Choosing $u_n = U e_n, n \geq 1$,

$$x_n^2 = K^{-1}[P + \sum_{i=1}^K \lambda_i - K\lambda_n], \quad n=1, \dots, K$$

$$x_n^2 = 0, \quad n > K$$

one obtains (a) immediately if H_2 is finite-dimensional. If H_2 is infinite-dimensional, then (a) is obtained as above by taking M sufficiently large, and noting that $\sum_1^M \log[1 + \tau_n]$ is a non-decreasing function of M if $\tau_n \geq 0$ for $n \geq 1$. The proof of (a) under the additional constraint that support $(\mu_X \circ A^{-1}) \leq M < \infty$ can be obtained from the preceding proof and the results of [1, pp. 83-84].

Suppose now that $\sum_1^K \lambda_n + P > K\lambda_K$ for all $K \geq 1$. If $P < \sum_1^\infty |\lambda_n|$, then there exists $K \geq 1$ and $\Delta > 0$ such that $P + \sum_1^K \lambda_n = -\Delta$. Thus,

$$|\lambda_{K+1}| > (1/K) [-\sum_1^K \lambda_n - P] = \Delta/K.$$

Assume $|\lambda_{K+p}| > (1/K)\Delta$ for $1 \leq p \leq N$.

Then

$$|\lambda_{K+N+1}| > \frac{1}{K+N} \left[-\sum_1^{K+N} \lambda_n - P \right] = \frac{1}{K+N} \left[\Delta - \sum_{K+1}^{K+N} \lambda_n \right]$$

so that

$$|\lambda_{K+N+1}| > \frac{1}{K+N} [\Delta + (N/K)\Delta] = \Delta/K.$$

Hence, $|\lambda_{K+p}| > \Delta/K$ for all $p \geq 1$, which contradicts $\lambda_n \rightarrow 0$, proving

$$\sum_1^\infty |\lambda_n| \leq P.$$

A lower bound on C under the assumptions of (b) is now obtained from the proof of (a), by taking C_K to be the value of C under the additional constraint that support $(\mu_X \circ A^{-1})$ has dimension $\leq K$, so that

$$C_K = \frac{1}{2} \sum_{n=1}^K \log \left[\frac{\sum_1^K \lambda_i + P}{K(1+\lambda_n)} \right]$$

and so

$$C \geq \lim_K C_K = \frac{1}{2} \sum_1^\infty \log[(1+\lambda_n)^{-1}] + \frac{1}{2}(P + \sum_1^\infty \lambda_n).$$

To see that $\lim_K C_K$ is an upper bound for C , suppose that $\lim_K C_K < C$. Then there exists a Gaussian $\mu_X \circ A^{-1}$ with covariance operator

$$R_{A(X)} = \sum_1^\infty \tau_n' [R_N^{1/2} u_n'] \circ [R_N^{1/2} u_n'] \quad \text{with}$$

$$\sum_1^\infty \tau_n' ||(I+S)^{1/2} u_n||_2^2 \leq P, \text{ and for some finite } M$$

$$\frac{1}{2} \sum_1^M \log[1 + \tau'_n] > \lim_K C_K.$$

However, as already seen, no selection of $\{(\tau'_n), (u'_n)\}$ satisfying the above conditions can be such that

$$\frac{1}{2} \sum_1^M \log[1 + \tau'_n] > C_M.$$

This contradiction establishes (b). The fact that the capacity in (b) cannot be attained follows in the same way. \square

Lemma 4 Let $C_{(N)} \equiv \sup_{\{(A, \mu_X): E_{\mu_X} \|R_N^{-1/2} A(X)\|_2^2 \leq P\}} I[\mu_{XY}]$.

(a) If S has only non-negative eigenvalues, then $C \leq C_{(N)}$.

(b) If S has only non-positive eigenvalues, then $C \geq C_{(N)}$.

(c) If H_2 is infinite-dimensional and all eigenvalues of S are non-negative, then $C = C_{(N)} = P/2$.

Proof (a) Here $\|(I+S)^{-1}\| \leq 1$, and so $\|R_N^{-1/2} A(X)\|_2^2$

$$\leq \|(I+S)^{-1}\| \|R_W^{-1/2} A(X)\|_2^2 \leq \|R_W^{-1/2} A(X)\|_2^2$$

for all $A(X)$ in range $(R_N^{1/2})$. Hence $E_{\mu_X} \|R_W^{-1/2} A(X)\|_2^2 \leq P$ implies

$E_{\mu_X} \|R_N^{-1/2} A(X)\|_2^2 \leq P$, so that C is obtained by a supremum over a smaller set, yielding (a).

(b) In this case $\|I+S\| \leq 1$, and so

$$E_{\mu_X} \|R_N^{-1/2} A(X)\|_2^2 \leq P \text{ implies } E_{\mu_X} \|R_W^{-1/2} A(X)\|_2^2 \leq P.$$

(c) From (a) and [1, Theorem 2], $C \leq P/2$. For fixed $n \geq 1$, let $\mu_{A_n(X_n)}$ be

Gaussian with covariance $R_{A_n(X_n)} = \frac{C}{n} \sum_1^n [R_N^{1/2} u_i] \bullet R_N^{1/2} u_i$ with $\{u_i, i \geq 1\}$

a CON set in H_2 . From [1, Theorem 1], $I[\mu_{X_n Y_n}] = (n/2) \log[1 + \frac{C_n}{n}]$. Define (C_n) by

$$C_n = nP \left[\sum_{i=1}^n \left\| (I+S)^{\frac{1}{2}} u_i \right\|_2^2 \right]^{-1}. \text{ Then } E_{\mu_{X_n}} \left\| R_W^{-1/2} A(X_n) \right\| = P, \text{ so } (A, \mu_{X_n}) \in Q.$$

The fact that $(1/n) \sum_{i=1}^n \left\| (I+S)^{\frac{1}{2}} u_i \right\|_2^2 \rightarrow 1$ follows easily from the fact that

$$\left\| S^{1/2} u_i \right\|_2^2 \rightarrow 0. \text{ Thus, } C_n \rightarrow P, \text{ and so } I[\mu_{X_n Y_n}] \rightarrow P/2, \text{ showing } C \geq P/2. \quad \square$$

Proof of Theorem

(a) is proved exactly as (a) of Lemma 3. (d) is contained in Lemma 4.

To prove (b) and (c) we identify two possibilities for each: (1) S has a finite set of M strictly negative eigenvalues; (2) S has an infinite set of strictly negative eigenvalues. We prove only (b) and (c) for case (1); the proof for (2) is similar but simpler.

First, we note that the capacity is at least C , as given in (b) and (c).

This follows immediately in (c) using a Gaussian μ_{AX} with covariance as specified for (c) in the theorem. For (b), this is shown by using a sequence of Gaussian signals (μ_{AX}^k) having covariance

$$R_{AX}^k = - \sum_{i=1}^M \frac{\lambda_i}{1+\lambda_i} R_N^{1/2} e_i \otimes R_N^{1/2} e_i + \frac{P + \sum_{i=1}^M \lambda_i}{k-M-1} \sum_{i=M+1}^k R_N^{1/2} u_i \otimes R_N^{1/2} u_i, \text{ where}$$

$S e_i = \lambda_i e_i$, $\lambda_i < 0$, $S u_j = \gamma_j u_j$, $\gamma_j \geq 0$, for $i=1, \dots, M$; $j=M+1, \dots, k$; and $\{e_1, \dots, e_M, u_{M+1}, \dots, u_k\}$ is an ON set. As $k \rightarrow \infty$, $I[\mu_{AX}^k] \rightarrow \frac{1}{2} \sum_{i=1}^M \log \left[\frac{1}{1+\lambda_i} \right] + \frac{1}{2} (P + \sum_{i=1}^M \lambda_i) = C$.

Thus, to prove (b) and (c) when S has only a finite set of strictly negative eigenvalues, it suffices to show that the capacity is $\leq C$ as given in the theorem.

Suppose then that the assumptions of either (b) or (c) are satisfied, and that S has strictly negative eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_M$. Suppose that the true capacity is greater than C , as given by (b) or (c). Since the capacity is the limit of

$(I[\mu_{XY}^k])$ for a sequence $(\mu_{X^k} \circ A_k^{-1})$ of Gaussian measures, there exists [1] a covariance operator R such that $R = \sum_1^J \tau_j R_N^{1/2} u_j \otimes R_N^{1/2} u_j$, $\frac{1}{2} \sum_1^J \log[1 + \tau_j] > C$, $\{u_j, j \geq 1\}$ an o.n. set, and $\sum_1^J \tau_j \|(I+S)^{\frac{1}{2}} U^* u_j\|_2^2 = P_1 \leq P$. We can assume that $J > M$. Let $T: H_2 \rightarrow H_2$ be the unitary map defined by $T v_j = U^* u_j$, where $\{v_j, j \geq 1\}$ are the eigenvectors of S . Defining $x_j^2 = \tau_j \|(I+S)^{\frac{1}{2}} U^* u_j\|_2^2$, one has that

$$\begin{aligned} \sum_1^J \log[1 + \tau_j] &\leq \sum_1^J \log[1 + x_j^2 \|(I+S)^{-\frac{1}{2}} U^* u_j\|_2^2] \\ &= \sum_1^J \log[1 + x_j^2 \|(I+S)^{-\frac{1}{2}} U^* T v_j\|_2^2]. \end{aligned}$$

We can assume that $x_j^2 \geq x_{j+1}^2$, $j=1, \dots, J-1$; from Lemma 2,

$$\begin{aligned} \sum_1^J \log[1 + \tau_j] &\leq \sum_1^M \log[1 + x_j^2 (1 + \lambda_j)^{-1}] \\ &\quad + \sum_{M+1}^J \log[1 + x_j^2 \|(I+S)^{-1/2} T v_j\|_2^2]. \end{aligned}$$

Define $P_1' = \sum_1^M x_j^2$. Then $\sum_{M+1}^J x_j^2 = P_1 - P_1'$; the eigenvalues of $(I+S)^{-1}$ are the same as those of $T^*(I+S)^{-1}T$, and thus Lemma 3 and Lemma 4(a) yield

$$\sum_1^J \log[1 + \tau_j] \leq C_0(P_1') \equiv \sum_{n=1}^K \log \left[\frac{\sum_1^K (1 + \lambda_i) + P_1'}{K(1 + \lambda_n)} \right] + P_1 - P_1'.$$

We maximize $C_0(P_1')$ with respect to P_1' . The derivative is non-decreasing for increasing P_1' if $\sum_1^K \lambda_i + P_1' \leq 0$. This is satisfied if $K < M$; taking P_1 arbitrarily close to P ($J \rightarrow \infty$), one obtains part (c) of the theorem. If $K = M$, then

$\frac{d}{dP_1'} C_0(P_1') > 0$ if $P_1' < -\sum_1^M \lambda_i$, $\frac{d}{dP_1'} C_0(P_1') < 0$ if $P_1' > -\sum_1^M \lambda_i$, and by continuity of C_0 , $C_0(P_1') \leq C_0(-\sum_1^M \lambda_i)$, proving (b) of the theorem.

The remaining parts of the theorem can be obtained from the proof of Lemma 3, and from Lemma 4. □

References

- [1] C. R. Baker, Capacity of the Gaussian channel without feedback, *Information and Control*, 37, 70-89 (1978).
- [2] P. M. Ebert, The capacity of the Gaussian channel with feedback, *Bell System Tech. J.*, 49, 1705-1712 (1970).
- [3] R. M. Fano, Transmission of information, MIT Press, Cambridge, MA and Wiley, New York (1961).
- [4] R. G. Gallager, Information theory and reliable communication, Wiley, New York (1968).
- [5] A. Horn, On the singular values of a product of completely continuous operators, *Proc. Nat. Acad. Sci.*, 36, 374-375 (1950).
- [6] S. Ihara, On the capacity of the continuous time Gaussian channel with feedback, *J. Multivariate Analysis*, 10, 319-331 (1980).
- [7] J. Stoer and C. Witzgall, Convexity and optimization in finite dimensions I, Springer-Verlag, Berlin (1970).

UNCLASSIFIED

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER C	GOVT ACCESSION NO. AD A105659	RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Information Capacity of Gaussian Channels.	5. TYPE OF REPORT & PERIOD COVERED TECHNICAL r.t.t.	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Charles R. Baker	8. CONTRACT OR GRANT NUMBER(s) ONR Contract N00014-75C-0491	9. PERFORMING ORGANIZATION NAME AND ADDRESS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program (Code 436) Arlington, VA 22217	12. REPORT DATE September 1981	10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 27	13. NUMBER OF PAGES 18	15. SECURITY CLASS. (of this report) UNCLASSIFIED
16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release -- Distribution Unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Consider a communication channel without feedback, with transmitted signal $A(X)$, and with additive Gaussian noise N . The information capacity of this channel is obtained subject to the constraints $E A(X) _W^2 \leq P$, where $ \cdot _W$ can be regarded as the RKHS norm of the stochastic process W . The class of admissible processes W includes all Gaussian processes having induced measure		

UNCLASSIFIED

182850

UNCLASSIFIED

20. equivalent to the measure induced by N.

UNCLASSIFIED

ATE
LMED
— 8